Identification of novel SNPs (alleles) associated with resistance to spot blotch in Wheat (*T. aestivum* L.)

A Thesis Submitted to

Institute of Advanced Research
Koba Institutional Area, Gandhinagar-382426
Gujarat, India

Thesis
SUBMITTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
BIOTECHNOLOGY

SUBMITTED
BY
VIPIN TOMAR
(Reg. No UIAR/11078)

Supervisor
Budhi Sagar Tiwari, PhD
Associate Professor
Plant molecular biology laboratory
Department of biological sciences & biotechnology
Institute of Advanced Research
Gandhinagar, Gujarat, India

Co-Supervisor
Uttam Kumar, PhD
Scientist (Wheat Breeder)
CIMMYT/Borlaug Institute for South Asia, India

March 2021

© Copyright by VIPIN TOMAR, 2021 All Rights Reserved

LIST OF CONTENTS

S. No.	TITLES	Page No.
	LIST OF FIGURES AND TABLES	
	ABBREVIATIONS	
	ABSTRACT	
	INTRODUCTION	1
1.0	Importance of cereal crops	3
1.1	Genome-wide association studies in spot blotch resistance related to bread wheat (*T. aestivum* L.)	5
	REVIEW OF LITRATURE	8
2.0	Origin and domestication of wheat	8
2.1	Spot blotch disease	11
2.2	Pathogen variability	13
2.3	*Bipolaris sorokiniana* (Sacc.) and its life cycle	13
2.4	Symptoms and diagnosis of the disease	15
2.5	Disease assessment	15
2.6	Genome-wide selection in wheat breeding	16
2.7	Components in genomic selection scheme	18
2.8	Factors affecting the genomic prediction accuracy	20
2.8.1	Size of training data	20
2.8.2	Relationship between the training population and test individuals	21
2.8.3	Cross-validation	21
2.8.4	Marker density	23
2.8.5	Prediction accuracy of statistical models	23
2.9	Implementation of genomic selection	25
2.10	Genotyping-by-sequencing (GBS)	26
2.11	Application of GBS in plant breeding	27
2.12	Previous studies on spot blotch resistance related to bread wheat (*T. aestivum* L.)	28
	MATERIALS AND METHODS	30
3.0	Plant genetic material	30
3.1	Genotyping-By-Sequencing and SNP filtering	30
3.2	Genetic analysis of SNPs	30
3.3	Population Structure Analysis	30
3.4	Genetic diversity analysis and analysis of molecular variance (AMOVA)	31
3.5	Plant materials and field layout	32
3.6	Screening for spot blotch disease resistance	33
3.7	Genotyping	33
3.8	Statistical analysis	33
3.9	Principal component analysis and linkage disequilibrium	34
3.10	Genome-wide association analysis	35
3.11	Gene functional annotations	35
3.12	Physical Mapping	35
3.13	Plant materials	36
3.14	Field trials and phenotypic evaluation	36

3.15	Genotyping-by-sequencing and SNP filtering	36
3.16	Statistical analysis of phenotypes	36
3.17	Baseline single-environment (SE) GBLUP model, Cross-validation schemes and predictive ability	37
3.18	Advanced multi-environment (ME) GBLUP model, Cross-validation schemes and predictive ability	38
	RESULTS (Chapter- I) **Elucidating SNP-based genetic diversity and population structure of advanced breeding lines of bread wheat (*Triticum aestivum* L.)**	40
4.0	SNPs distribution on the wheat genome	40
4.1	Genetic analysis of SNPs	41
4.2	Population structure analysis	42
4.3	Analysis of Molecular Variance (AMOVA)	44
4.4	Allelic pattern covering the populations	44
	RESULTS (Chapter- II) **New QTLs for spot blotch disease resistance in wheat (*Triticum aestivum L.*) using genome-wide association mapping**	47
4.5	Estimation of heritability	47
4.6	SNP density and principal component analysis	48
4.7	Marker-trait associations of spot blotch	50
4.8	Physical mapping	55
4.9	Putative candidate genes and annotations	56
	RESULTS (Chapter- III) **Increased predictive accuracy of multi-environment genomic prediction model for yield and related traits in spring wheat (*Triticum aestivum* L.)**	63
4.10	Heritability, correlations and traits characterization	63
4.11	Baseline SE model: Performance of untested lines in the same environment	68
4.12	Advanced ME model: Performance of tested lines in untested environments and untested in tested environments	69
5.0	**DISCUSSION**	70
5.1	Phenotypic evaluation for spot blotch	74
5.2	Principal component analysis	75
5.3	SNP effects in different environments of spot blotch	75
5.4	Trait correlation and characterization (A vital factor for improving accuracy in ME-GP)	80
5.5	Yield and related traits heritability difference among environments	80
5.6	SE and ME genomic prediction across years and sites and its utilities in crop breeding	81
6.0	**CONCLUSIONS**	83
7.0	**FUTURE PRESPECTIVES**	85
	SUMMERY	86
	REFERENCES	89

	SUPPLEMENTARY FILES	
	Publications information (Research article copy)	
	Attended Conferences/Workshops	

ABSTRACT

Genetic diversity and population structure information are crucial for enhancing traits of interest and the development of superlative varieties for commercialization. The present study elucidated the genetic diversity and population structure of 141 advanced wheat breeding lines using single nucleotide polymorphism (SNP) markers. A total of 14563 high-quality identified genotyping-by-sequencing (GBS) SNPs were distributed across the wheat genome covering 13.9 GB, with a minimum of 1026 SNPs on homoeologous group four and maximum of 2838 SNPs on group seven. The average minor allele frequency was found 0.233, while the average polymorphism information content (PIC) and heterozygosity were 0.201 and 0.015, respectively. Principal component analyses (PCA) and population structure identified two major groups (sub-populations) based on SNPs information. The results indicated a substantial gene flow/exchange with many migrants (Nm = 86.428) and a considerable genetic diversity (number of different alleles, Na = 1.977; the number of effective alleles, Ne=1.519; and Shannon's information index, *I* = 0.477) within the population, illustrating a good source for wheat improvement. The average PIC of 0.201 demonstrates moderate genetic diversity of the present evaluated advanced breeding panel. Analysis of molecular variance (AMOVA) detected 1% and 99% variance between and within subgroups. It is indicative of excessive gene traffic (less genetic differentiation) among the populations. These conclusions deliver important information with the potential to contribute new beneficial alleles using genomic selection and genome-wide association studies (GWAS) to enhance genetic gain in South Asian wheat breeding programs.

INTRODUCTION

Wheat is a self-pollinating plant belonging to the family *Poaceae* (grasses) and genus *Triticum*. Wheat is called "the plant of all plants" Since it is the oldest plant cultivated by humankind and domesticated ten thousand years ago (Charmet, 2011; Marcussen et al., 2014). Bread wheat (*Triticum aestivum* L.) is an allohexaploid species originating from successive rounds of hybridization in the Fertile Crescent during the Neolithic time, ~8000 to 10,000 years ago (Sansaloni et al., 2020). Wheat has become a vital cereal crop of the world (after rice and maize), establishing an essential food source to a vast population, particularly in Asia. Wheat has become a staple food for the human population. It also provides a vital nutritious feed to animals with wheat straw being fed as a complement to green fodder. Wheat is a source of Carbohydrate, protein, minerals, fat, and fiber. It fortifies almost half of the world's energy supplies (Shewry, 2007; Topping, 2007; Simmonds, 2019). Globally more than eight billion tons of wheat is produced annually around the world. China, India, USA, France, Russia, Canada, Australia, Pakistan, Turkey, and Italy are the major wheat-producing countries that contribute about one-third of the total world wheat production. From 2014 to 2024, the demand for wheat would grow gradually in the Southeast Asian countries (Taylor D & Koo, 2015). India has increased wheat production by 120% since the 1980s since the high yielding varieties were introduced. The total area is about 30 million hectares of wheat crop in the country. There was a record production of more than 100 million tons of wheat in 2019-20.

Wheat (*Triticum aestivum* L.) is the primary staple food for more than 35% of the world’s population (FAO, 2018). Wheat is among the most consumed grain worldwide, providing 15% of calories every day, covering an area of more than 220 million hectares with almost 750 metric megatons production every year (Balfourier et al., 2019). Trends in population growth, food consumptions, and economic and socio-economic status in Africa and Asia continents have effected global wheat demand over the past few years (Mondal et al., 2016a). The pace of wheat improvement must accelerate to meet the projected global food demand by 2050. The Green revolution played a crucial role in India, Pakistan, Nepal, and Bangladesh, ensuring food security in this world's densely

populated region (Joshi et al., 2007b). However, wheat production faces multiple threats via rapidly evolving pathogen variants, pests, and increased climate variability, which considerably jeopardizes crop productivity (Sharma, Duveiller & Ortiz-Ferrara, 2007a; Friesen, Xu & Harris, 2008; Gurung et al., 2009, 2012). Breeding wheat for climatic resilience and disease resistance combined with good agronomic value can improve wheat productivity to meet future food demands (Mondal et al., 2016a).

The first generation of morphological markers could not present the actual picture of diversity because of their limited number, lack of information about environmental and epistatic interaction. Slowly the paradigm shifted to investigate genetic diversity and population structure using PCR (polymerase chain reaction) based markers, which gave a better view of underlying diversity markers due to their abundance and environmentally neutral nature (Farahani et al., 2019). Among various molecular markers the SNP markers are preferred molecular markers in several crops due to extensive genome coverage, chromosome-specific location, low cost, co-dominant inheritance, and fast-tracking compared to other PCR based markers. SNP are evolutionarily stable due to lower rates of recurrent mutation. SNPs are considered first-rate to understand genomic evolution and to study complex genetic traits. SNPs have been extensively used in genetic resource characterization, genome-wide association studies (GWAS), marker-assisted selection (MAS), and genomic selection (GS) (Ganal, Altmann & Röder, 2009).

Microarrays have been used as a pre-eminent answer to develop SNPs in polyploid wheat genomes (Wang et al., 2014). Once an all-inclusive SNP data is accessible for species, a cost-effective microarray may be formed, and the process is relatively convenient. Microarray avoids the risk of miscalling diversity on homoeologous genomes and recently with amplified 100- fold in wheat moving from 9 K (Cavanagh et al., 2013) to 820 K (Winfield et al., 2016). The 90 K wheat SNP array (Wang et al., 2014) has been effectively utilized for genetic diversity investigation, the building of high-density maps, and GWAS (Maccaferri et al., 2015a; Chen et al., 2016; Mengistu et al., 2016; Wen et al., 2017). The low cost of GBS makes it a robust method for detecting and genotyping SNPs in various crops. GBS has been efficiently applied in genomic

diversity, genetic linkage studies, genome-wide association study (GWAS), and genomic selection under large scale plant breeding programs. GBS is found to be a perfect platform from single-gene to whole-genome sequencing and suited to genetic analysis and marker development (Heffner, Sorrells & Jannink, 2009; Heffner et al., 2010; Jannink, Lorenz & Iwata, 2010; Fu & Peterson, 2011; Poland & Rife, 2012; Poland et al., 2012a; Thomson et al., 2012; Lu et al., 2013; Fu, Cheng & Peterson, 2014). The GBS has been established with maize and barley inbred populations with approximately 200,000 and 25,000 sequence tags (Elshire et al., 2011).

Population structure, genetic diversity, and relationships among breeding lines are vital in scheming appropriate breeding plans (Peterson et al., 2014; Singh, Agarwal & Yadav, 2019). The population structure information aids in estimating the accurate association between phenotypic and genotypic variation (Knowler et al., 1988). This information allows researchers to utilize natural diversity for detecting vital genes/QTLs in the genome by using current genetic technologies for breeding (Zhu et al., 2008). Genetic diversity investigating and detection has advanced from the distinct morphological traits to DNA variation's molecular investigations (Zhang, Maroof & Kleinhofs, 1993). Determining the current genetic diversity of crops has paramount importance for selecting and conserving parents with the various genetic framework, thus providing well-organized crop enhancement (Mengistu, Kiros & Pè, 2015).

1.0 Importance of cereal crops

After maize, rice, wheat, and barley are the foremost contributors to cereal crop production across the world (FAOSTAT., 2018). The domestication of these crops is dated back to about 10,000 years (Garris et al., 2005; Sang, 2009). Maize is majorly cultivated in the USA, Brazil, China and European countries while rice (*Oryza sativa*) is cultivated predominantly in Asian and African countries. Primary producers of wheat (*Triticum spp.*) are China, India, USA, and Europe. In contrastingly, the most important barley (*Hordeum vulgare*) producers are Russia, European countries and Canada (FAOSTAT., 2018).

Interestingly, the cultivation of wheat and rice is mostly to fulfil the needs of human nutrition. In contrast, maize and barley largely contribute as a malt

source in brewing industries and animal feed (Oerke & Dehne, 2004) or as a raw material for industrial production of starch-based oil and ethanol (Chaudhary, Kumar & Langyan, 2013). To meet the global world population demand, it is estimated that production in the agricultural sector must double by 2050 (Tilman et al., 2011; Gerland et al., 2014). However, climate changes and the loss of agricultural land obstruct production (Godfray et al., 2010). Under these circumstances, the present yield trend would possibly be insufficient to reach the foretold demand for agricultural production until 2050 (Ray et al., 2013). It has been observed that about 10 to 16% of yield loss of cereal crops is due to fungal pathogens (Oerke, 2006; Chakraborty & Newton, 2011). Pathogens often develop resistance against fungicides pesticides (Borel, 2017). Also use of pesticides is always questionable because of environmental problems over the need for higher production (Lamichhane et al., 2016). It motivates us to adopt new agricultural practices to overcome disease-caused yield losses.

More than three decades before, Saari (1985) documented that spot blotch is caused by the fungal pathogen *Bipolaris sorokiniana* (Sacc.) Shoem. Spot blotch is one of the main constraints to wheat in warmer and humid cropping areas, particularly in the Eastern Gangetic Plains (EGPs) of South Asia. In India, spot blotch is of much economic importance in the North Eastern plains zone (Chatrath et al., 2007; Joshi et al., 2007a), but due to climate change and crop intensification, spot blotch is capturing attention in the northwestern plains zone, and in the central zone of India, posturing new tasks to researchers (Chand et al., 2003; Joshi et al., 2007b). The occurrence of *Bipolaris sorokiniana* as a wheat pathogen in the North-Western part of the Russian Federation (Smurova, 2008) proposes that spot blotch potentially could become a serious threat to Europe and other western countries, as seen in recent studies (Adhikari et al., 2012; Duveiller & Sharma, 2012; Gurung et al., 2014; Ayana et al., 2018).

Due to spot blotch disease, yield losses range from 18-50% (Duveiller et al., 2005; Saari, 1998; Singh et al., 2004). The disease initially starts as brown eye-shaped necrotic spots fenced by a yellow halo on leaves. Later on, such spots merge to make a larger leaf area blighted and thus adversely affect

photosynthesis. These symptoms induced by *Bipolaris sorokiniana* are observable on leaves, sheath and stem (Chand et al., 2003). The crop becomes additional susceptible, especially after anthesis and at a later advanced plant development stage (Duveiller et al., 2005). The combined effects of high temperature, humidity, and long periods (>12 h) of leaf wetness caused by rainfall or dew are conducive to foliar blight development (Duveiller, 2004). Approximately 25 million hectares of wheat cultivated land is affected by spot blotch disease worldwide (van Ginkel & Rajaram, 1998), and around 11 million hectares is located in India alone (Joshi et al., 2007c).

Resistance to spot blotch in wheat is a quantitative trait (Joshi et al., 2004). It is, therefore, necessary to identify the quantitative trait loci (QTLs) conferring disease resistance. Quantitative traits genetic analysis has been revolutionized by the introduction of environmentally insensitive DNA-based molecular markers. The DNA-based markers are not influenced by environmental factors and/or the plant developmental stage. They, therefore, represent a promising, environment-insensitive means for selecting genotypes with improved spot blotch resistance. QTL mapping requires the collection of genotypic (molecular marker) and phenotypic data from a segregating population, followed by statistical analysis to reveal all possible genetic loci where allelic variation correlates with the phenotype.

1.1 Genome-wide association studies in spot blotch resistance related to bread wheat (*T. aestivum* L.)

Spot blotch caused by *Bipolaris sorokiniana* is a major constraint in wheat production in tropics and subtropics (Dubin & Van Ginkel, 1991; van Ginkel & Rajaram, 1998). The pathogen has a worldwide dispersal, but it is predominantly aggressive under warm, high relative humidity, and high temperature associated with imbalanced soil fertility. Major Yield losses are observed in fields with lower inputs and under late-sown conditions (Joshi et al., 2007b). *Bipolaris sorokiniana* act as a causal agent for numerous diseases like head blight, seedling blight, foliar blight/spot blotch, common root rot, and black point of wheat, barley, other small cereal grains and grasses (Lennard, 1984). Therefore, spot blotch of wheat is considered one of the most crucial

diseases caused by this pathogen in the mega-environments characterized by temperature above 17˚C and high humidity (van Ginkel & Rajaram, 1998).

It is often difficult to achieve the desired level of resistance to several diseases through conventional breeding. In the case of wheat disease, resistance is inherited qualitatively and quantitatively (Kumar et al., 2009, 2010; Marone et al., 2013; French, Kim & Iyer-Pascuzzi, 2016). The genetic basis of spot blotch resistance has earlier been documented as eight major quantitative trait loci (QTLs), namely *QSb.bhu-2A, QSb.bhu-2B, QSb.bhu-2D, QSb.bhu-3B, QSb.bhu-5B, QSb.bhu-6D, QSb.bhu-7B* and *QSb.bhu-7D* (Kumar et al., 2009, 2010). (Sharma, Duveiller, & Jacquemin, 2007b) reported three microsatellite markers (*Xgwm67*, *Xgwm570,* and *Xgwm469)* linked with spot blotch resistance. The QTL *QSb.bhu-5B, QSb.bhu-7D, and QSb.bhu-3B* have been designated as genes *Sb1*, *Sb2, Sb3 and Sb4* respectively in further studies along with Sb4 at 4B (Lillemo et al., 2013; Kumar et al., 2015; Lu et al., 2016; Zhang et al., 2020). *Lr34* and *Lr46*, the two broadly used genes conferring leaf rust resistance, have also been reported to contribute to spot blotch resistance. While the *Lr34* gene has been reported to be the major locus on chromosome 7D explaining up to 55% phenotypic variation for spot blotch disease resistance, this locus was given the gene designation *Sb1* (Lillemo et al., 2013). Nevertheless, in the past few years, several QTLs and genetic markers for spot blotch resistance have been identified in multiple studies in wheat (Gurung et al., 2014; Zhu et al., 2014; Singh et al., 2018)

Due to evolutionary changes in pathogen populations, resistance genes may lose their effectiveness over time. Given these challenges, identification and mapping of novel resistance genes would aid breeding for disease resistance in wheat. One of the approaches is to identify spot blotch resistance QTLs through association mapping. This approach leverages historical recombination and generally has better mapping resolution compared to biparental mapping (Zhu et al., 2008). The genome-wide association study (GWAS) using genotyping-by-sequencing (GBS) to identify QTLs for traits of interest provides advantage through better genome coverage compared to conventional marker systems like SSRs, AFLP, CAPs, etc. The GBS utilizes the advantage of high throughput genotyping assays with relatively low data costs, which are much

higher in genome sequencing and re-sequencing techniques (Elshire et al., 2011). A fundamental approach in GWAS is to have enough genome coverage so that functional alleles will be in linkage disequilibrium (LD) with at least one marker (Myles et al., 2009). The association studies for disease resistance, including spot blotch, have been reported in some of earlier studies (Maccaferri et al., 2010, 2015b; Yu et al., 2011; Zegeye et al., 2014; Ayana et al., 2018).

There are limited reports where the same set of genotypes are exposed to natural disease pressure to identify genomic regions underpinning wheat spot blotch resistance. Thus, the main objective of this study was to establish marker-trait associations for spot blotch resistance using GBS markers in spring wheat specific to the South Asian regions, viz. India and Bangladesh, the well-known hot spots for this disease. Therefore, the present study was proposed to identify novel QTLs and validated known genomic loci conferring spot blotch resistance in 141 advanced wheat breeding lines phenotyped under natural field conditions at Borlaug Institute for South Asia, PUSA, Bihar, and BARI, Bangladesh.

Objectives

- To evaluate the advanced wheat breeding genotypes for resistance against spot blotch caused by *Bipolaris sorokiniana.*
- To evaluate the genetic diversity and population structure within this CIMMYT wheat advanced breeding lines.
- To identify/locate QTLs of spot blotch using genome-wide association study. Identification of markers associated with the resistance loci will facilitate the marker-assisted selection of spot blotch resistance in the wheat breeding program.

6.0 CONCLUSIONS

In the present study, we applied GBS based SNP to learn GBS-SNP markers' usefulness for diversity analysis in 141 elite wheat breeding lines. Despite very designated, our advanced breeding lines panel found to be genetically diverse, which could be instrumental for future South Asian breeding programs to develop new elite wheat varieties of alluring traits i.e., high yield, biotic and abiotic resistance. Besides, the present study dappled two subpopulations that were enlightened by their parentage and selection history. The low heterozygosity detected among elite advanced wheat breeding lines within subpopulations and the moderate divergence among sub-populations suggested that the elite advanced wheat breeding lines could be used further for GWAS and genomic selection studies.

Furthermore, we identified genetic regions underlying spot blotch resistance in the elite spring wheat genotypes. The variable conditions at three field environments in India and Bangladesh allowed us to capture the considerable phenotypic variation for spot blotch disease with the GWAS, resulting in a total of eight QTLs belonging to eight wheat chromosomes. The literature mining of the functional gene annotations of identified SNPs encoding the single protein or protein family directly or indirectly involved in disease resistance has led to the identification of putative target genes and functions to determine the disease-resistance mechanism. The new QTLs appeared on chromosomes 1A (497.2), 1D (89.84), 2B (421.92), and 6D (6.84) associated with many disease resistance family proteins. The SNP on chromosome 2A was found to be associated with a known gene that encodes ‘senescence-associated family protein’ and is directly involved in spot blotch resistance. The results are of importance for the breeders in developing spot blotch resistant varieties targeting the South Asian wheat-growing regions. Given the aggressive pathogen spread and food security concerns, the breeding programs in South Asia could benefit from the present study. The mapping of favorable alleles can facilitate introgression of the alleles into present-day elite cultivars to impart disease resistance. Additional investigations are underway, which would further confirm the importance of these chromosomal regions/genes associated with spot blotch.

Furthermore, escalating breeding for quantitative traits which is challenging due to the complex genetic architecture of traits that are highly affected by the complex G x E interactions in the field trials. A suitable genomic prediction modelling strategy can potentially address this challenge through multi-environmental (ME) genomic prediction models. We evaluated genomic prediction accuracies of advanced spring wheat lines under four diverse environments in two wheat-growing regions in India. ME genomic selection models showed significant improvement over single-environment (SE) models in terms of prediction accuracies. Our results suggest that ME can be leveraged to improve breeding selection efficiency for major agronomic and phenological traits. Over the years, CIMMYT has established an extensive network of field-testing sites in South Asian countries including India, Pakistan, Bangladesh, and Nepal. Our results suggest that the wheat breeding programs in these countries can greatly benefit from genomic selection through better modelling of environmental variance and sparse testing of a larger cohort of breeding lines. Future research efforts will be directed towards including high-throughput phenotyping traits such as plant height, NDVI, and senescence into the genomic prediction framework to improve the selection efficiency of spring wheat in the South Asian breeding programs.

7.0 FUTURE PRESPECTIVES

The present study elucidated that spot blotch response in wheat is controlled by many QTLs signifying the compound genetic control of spot blotch. After validation above-identified markers can be further utilized for MAS in developing an improved source of spot blotch resistance in wheat breeding programs in future. Several other genes and transcription factors are also known to be involved in spot blotch. Therefore, in future, an all-inclusive systematic study including all the genes/QTL/pathways involved in spot blotch in wheat may allow a better understanding of disease in wheat and also help in formulating a thorough approach for wheat breeding. Even a marginal reduction in disease level may be of significance for wheat-growing areas in developing countries. Thus, the information generated here will allow improvement of spot blotch resistant wheat through marker-assisted selection strategy in future South Asian wheat breeding programs.

SUMMARY

Bread wheat is one of the world's most important cereal food crop, but yields are constantly challenged by various biotic constrains. Spot blotch caused by *Bipolaris sorokiniana* is an economically important disease in many warm and humid wheat-growing areas like Eastern India, Nepal, China, South East Asia, Latin America, and sub-Saharan Africa because of its potential to cause considerable yield losses and reduced grain quality. In South Asia, the importance of wheat in household food security is well recognized. This disease is of economic importance at the national and international level and has a large impact on Indo-Gangetic Plains in India. Developing resistant cultivars using molecular breeding is likely to be a robust, economical and environmentally friendly means of controlling this disease. The present investigation was undertaken with the aim to phenotype the advanced breeding lines at different hot spot locations namely, BISA, Bihar, India and BARI, Bangladesh for three seasons (2016-17, 2018-19, and 2019-20) under natural field conditions and to identify new QTLs for spot blotch resistance in wheat.

A summary of the material and methods and the results are presented below.

1. The elite advanced wheat breeding lines comprising 141 lines imported from CIMMYT, Mexico were used for genome-wide association mapping to determine marker-trait associations (MTAs). The panel was evaluated for three years (2016-17 and 2018-19 and 2019-20) at two locations PUSA, Bihar and BARI, Bangladesh, in two replications independently environments.
2. Genotyping data of 14,563 SNP markers (developed through GBS), spread over all the 21 wheat chromosomes, were generated by Poland Lab (Kansas State University, USA under the FEED THE FUTURE project of USAID and this data were used for GWAS in the present study.
3. The present study elucidated the genetic diversity and population structure of 141 advanced wheat breeding lines using single nucleotide polymorphism (SNP) markers.
4. A total of 14563 high-quality identified genotyping-by-sequencing (GBS) SNPs were rationally distributed covering the wheat genome.

5. Minimum 1026 SNPs were in homoeologous group four while a maximum of 2838 SNPs were found in group seven. The lowest minor allele frequency was found 0.05, while the average polymorphism information content (PIC) and heterozygosity were 0.201 and 0.03, respectively.
6. Principal component analyses (PCA) and population structure identified two major groups (sub-populations) based on SNPs information. The results indicated a substantial gene flow/ exchange with many migrants (Nm = 38.299) and a considerable genetic diversity (number of different alleles, Na = 1.99; the number of effective alleles, Ne=1.512; and Shannon's information index, *I* = 0.482) within the population, illustrating a good source for wheat improvement.
7. The average PIC of 0.201 demonstrates moderate genetic diversity of the present evaluated advanced breeding panel. Analysis of molecular variance (AMOVA) detected 1% and 99% variance between and within sub-populations. It is indicative of excessive gene traffic (less genetic differentiation) among the populations.
8. The population were evaluated at BISA, PUSA India and BARI, Bangladesh sites over three crop seasons (2016-17, 2018-19 and 2019-20). The experimental fields at each site were managed under irrigated (IR) and rainfed (RF) conditions. Data were recorded on the following traits: Spot Blotch evaluation, days to heading (DTHD), days to maturity (DTMT), thousand-grain weight (TGW) and grain yield/plot (GRYLD).
9. Structure simulation-based cluster analysis identified two Major sub-groups (G1, G2) among 141 genotypes.
10. The estimates of CV (%) for spot blotch in different traits varied in different environments. The CV (%) were from 27.09 (E1), 24.70 (E2), and 61.59 % (E3), in different environments and 31.37 % in the combined environment. Taken together, the extent of available variability suggested the suitability of advanced breeding lines for the study of MTAs.
11. A total of 8 QTLs on associated with spot blotch disease resistance were identified across the genome on eight chromosomes. Among the 8 Functional annotations identified NBS-LRR, MADS-box transcription

factor, and 34 other plant-related protein families across multiple chromosomal regions associated with significant markers. FarmCPU algorithm identified marker-trait associations, and allele effects and candidate genes were further identified.

12. The physical map of the population comprising 29 MTAs was generated for the present study. The total SNP coverage was 13901.92 Mb with an average distance between SNPs was 1.048 Mb. The average SNPs density was 0.955 Mb.
13. The results indicate four promising new QTLs on chromosomes 1A (497.2 Mb), 1D (89.84 Mb), 2B (421.92 Mb), and 6D (6.84 Mb) associated with several disease resistance family proteins.
14. Breeding for quantitative traits is challenging due to the complex genetic architecture of traits that are highly affected by the complex G x E interactions in the field trials. Therefore, in this study, a suitable genomic prediction modelling strategy was developed to potentially address this challenge through multi-environmental (ME) genomic prediction models
15. We evaluated genomic prediction accuracies of advanced spring wheat lines under four diverse environments in two wheat-growing regions in India.
16. ME genomic selection models showed significant improvement over single-environment (SE) models in terms of prediction accuracies.
17. Our results suggest that ME can be leveraged to improve breeding selection efficiency for major agronomic and phenological traits.